IASOA Data & Observations Working Group

February 14, 2018

Attendees: Sara Morris, Taneil Uttal, Audra McClure, Barbara Casati, Chris Cox, Robin Stevens, Sandy Starkweather, Von Walden, Jonny Day, Matthew Shupe, Ola Persson

Role Call of group members

IASOA and YOPP:

- YOPP is the core component of the Polar Prediction Project (PPP), an international activity initiated by World Meteorological Organization's World Weather Research Program (WWRP) which spans 2013-2022
- YOPP will take place from mid-2017 go mid-2019 (want full annual cycle)
- YOPP goal: to significantly advance our environmental prediction capabilities for the polar regions and beyond
- A major project of the YOPP verification task team: supersites multivariate process-based evaluation (model focused)
- Special Observing Periods within YOPP were chosen for Arctic (Feb-March & July-August) and Antarctic (Feb-March 2019)
    - These periods are to concentrate focus on ensuring all obs available at during these periods (model output will remain the same during these SOPs as the rest of the year)
- Drafted list of observations details available/contributing from each of the IASOA stations online
- YOPP activities requiring observations/model validations from IASOA stations/observatories
- YOPP planning to archive model data around IASOA stations (sub-hourly frequency: a common timing which accommodates different model time steps will be defined; high-resolution models are expects to output a beam of grid-points surrounding the super-sites to enable representativeness studies)
- Observations require large effort to unify measurements and output formats
    - YOPP modeling community has not yet confirmed their file outputs
- ***IASOA AODF's: Amalgamated Observatory Data Files***
        - Will have standard format, to match model output files
        - Files can also be created for stations not within IASOA
        - Standard time series
        - Observation variables will be incorporated into one file
            - Per supersite
        - Observation specialists will determine most usable and representative values for variables that have multiple measurement outputs or derived techniques
        - Observation uncertainty is of interest for verification research studies: where multiple sources of obs are available would be valuable to have an assessment of the obs span. For observatories with significant topographic differences between measuring sites would be valuable to have multiple time series (as for the model output for a beam of gridpoints)

            - Additional non-YOPP variables will be included that supports observation science
        - Identify observation specialists via IASOA
        - Data flagging protocols
        - Initial focus will be on SOP's (special observing periods)
        - Publication of AODF data sets with doi and full author-list
        - Suggest that Antarctic supersites generate the same AODF's
    - YOPP: valuable to have any/all values available per measurement, don't have to decide on one specific value
    - YOPP: working on file name formatting and metadata formatting
        - Files will be classic netCDF, determining global attributes
        - CF conventions for variable names
        - Atomic files: single files for each model run (daily files –or twice a day- if more model runs are available)
            - Each file includes the vertical profile time-series (from the model run initial time through the model run lead time) for all variables and for all beam of grid-points surrounding an individual super-site
        - Will be difficult to identify single organization in naming convention for observations
        - Lat/lon information will also be included in metadata
        - Observatory files will have start and end times
            - Need to identify file time range for observation files (daily, etc.)
        - Model output will be station based – spatial or point, depends on model
            - Models have different array sizes, dimensions, parameters – so files will differ by model
    - Observation files: need to determine time series
        - Sub-hourly variability
        - 10 minutes could be nice for turbulent flux data
        - Need to get finalized YOPP data formatting document in order for observations to make decision on their files
        - AODF's will start with 15 minute time series
            - Seems like an optimal frequency since ECCC model time step is 90" and ECMWF model time-step is 5':15' the shortest common outputting frequency
        - Time convention: have both day fraction and individual date columns
            - Look into netCDF time format availability/standard

Action Items:

- Create google document to host station information and file format information (Uttal, ALL)
- Send around finalized YOPP data format document (Casati, Uttal)